

**CLAIMS:**

What is claimed is:

1. A method of producing a video signal at a set top box comprising:
  - receiving a first video signal at said set top box;
  - processing said first video signal to produce a first image stored in memory of said set top box;
  - receiving a second video signal at said set top box;
  - processing said second video signal to produce a second image stored in said memory of said set top box;
  - accessing a presentation description that defines a portion of said first image and that defines the manner in which said portion of said first image and a portion of said second image are combined;
  - combining said portion of said first image with said portion of second image in accordance with said presentation description to produce a combined image; and
  - displaying said combined image.
2. The method of claim 1 wherein said step of combining further comprises:
  - applying a mask that defines said portion of said first image.
3. The method of claim 1 wherein said step of combining said video signals further comprises:
  - generating a logical combination of said portion of said first image and said portion of said second image.
4. The method of claim 1 wherein said step of combining said video signals further comprises:
  - generating a mathematical combination of said portion of said first image and said portion of said second image.

5. The method of claim 1 wherein said step of combining said video signals further comprises:
  - scaling said portion of said first image.
6. The method of claim 1 wherein said step of combining said video signals further comprises:
  - warping said portion of said first image.
7. The method of claim 1 wherein said step of accessing said presentation description further comprises:
  - accessing said presentation description across a network.
8. The method of claim 1 wherein said step of accessing said presentation description further comprises:
  - receiving a network address at which a presentation description can be accessed.
9. The method of claim 1 wherein said step of accessing said presentation description further comprises:
  - selecting said presentation description from a plurality of presentation descriptions contained in said first video signal.  
5
10. The method of claim 1 further comprising:
  - modifying said presentation description in response to a user input.
11. The method of claim 1 further comprising:
  - processing said first video signal to produce first audio data stored in said memory of said set top box;
    - processing said second video signal to produce second audio data stored in said memory of said set top box;

5

accessing a presentation description that describes the manner in which said first audio data and said second audio data are combined; and combining said first audio data and said second audio data in accordance with said presentation description.

12. A method of displaying a sequence of combined images in a set top box comprising:

receiving a first video signal at said set top box;

processing said first video signal to produce a first sequence of images

5 stored in memory of said set top box;

receiving a second video signal at said set top box;

processing said second video signal to produce a second sequence of images stored in said memory of said set top box;

10 accessing a presentation description that defines a portion of said first sequence of images and that defines the manner in which said portion of said first sequence of images and a portion of said second sequence of images are combined;

15 combining said portion of said first sequence of images with said portion of said second sequence of images in accordance with said presentation description to produce a sequence of combined images; and

displaying said sequence of combined images.

13. The method of claim 12 wherein said step of combining further comprises:

applying a mask specified in said presentation description that defines said portion of said first sequence of images.

14. The method of claim 13 wherein said step of applying a mask further comprises:

executing program code that modifies said mask to select a different portion of at least one image of said first sequence of images.

15. The method of claim 12 wherein said step of combining said video signals further comprises:

generating a mathematical combination of said portion of one image of said first sequence of images and said portion of one image of said second sequence of images.

5  
16. The method of claim 12 wherein said step of combining said video signals further comprises:

generating a logical combination of said portion of one image of said first sequence of images and said portion of one image of said second sequence of images.

17. The method of claim 12 wherein said step of combining said video signals further comprises:

scaling said portion of one image of said first sequence of images .

18. The method of claim 12 wherein said step of combining said video signals further comprises:

warping said portion of one image of said first sequence of images.

19. The method of claim 12 further comprising:

modifying said presentation description in response to a user input.

20. A method of controlling generation of a combined video signal in a set top box unit at a user's premises from a broadcast site comprising:

transmitting a first digital video signal to said set top box;

transmitting a second digital video signal to said set top box substantially simultaneously with said first digital video signal;

loading image combination code into said set top box; and

providing a presentation description to said set top box that describes the manner in which a portion of an image contained in said first digital video signal

10       is combined with a portion of an image contained in said second digital video signal to produce said combined video signal.

21. The method of claim 20 wherein said step of providing a presentation description further comprises:

transmitting a network address that said set top box employs to access said presentation description..

22. The method of claim 20 wherein said step of providing a presentation description further comprises:

transmitting said presentation description to said set top box as a part of said first digital video signal.

23. The method of claim 20 wherein said step of providing a presentation description further comprises:

selecting said presentation description from a plurality of presentation descriptions wherein said presentation description conforms to the requirements of said set top box.

5       24. The method of claim 20 wherein said step of providing a presentation description further comprises:

altering a general presentation description to conform to the requirements of said set top box.

25. The method of claim 20 wherein said step of providing a presentation description further comprises:

tailoring a general presentation description to correspond to a viewer preference.

5       26. The method of claim 20 wherein said step of providing a presentation description further comprises:

transmitting a plurality of presentation descriptions to said set top box  
5 from which said set top box selects one presentation description that conforms to  
the requirements of said set top box.

27. A set top box that produces a combined video signal comprising:

a processor;

a memory;

5 a tuner/decoder that receives a first video signal and a second video signal substantially simultaneously and that routes control information contained in said first video signal to said processor and that routes first video data from said first video signal and second video data from said second video signal to a decoder;

10 said decoder that decodes said first video data and produces a first video image in said memory and that decodes said second video data and produces a second video image in said memory;

15 a presentation description stored in said memory that specifies the manner in which a portion of said first video image is combined with a portion of said second video image to produce said combined signal;

program code operating in said processor that employs said presentation description and that accesses said portion of said first video image and said portion of said second video image in said memory and that combines said first portion of said first video image and said portion of said second video image in a manner specified by said presentation description; and

20 a video output unit that outputs said combined signal to a display device.

28. The system of claim 27 further comprising:

a network interface that accesses a presentation description.

29. The system of claim 27 wherein said decoder further produces first audio data in said memory from said first video information and produces second audio data in said memory from said second video information.

30. The system of claim 29 wherein said presentation description further specifies the manner in which said first audio data is combined with said second audio data.

31. The system of claim 27 further comprising:

a user interface that receives an input from a user that modifies said presentation description.

32. The system of claim 27 further comprising:

user preference information stored in said memory that is used by said presentation description.

33. The system of claim 27 wherein said program code operating in said processor further comprises:

a software routine that controls said decoder to perform at least part of the combination of said portion of said first video image and said portion of said second video image in a manner specified by said presentation description.

34. The system of claim 27 wherein said program code operating in said processor further comprises:

a software routine that selects said presentation from a plurality of presentation descriptions contained in said first video signal.

35. A set top box that produces a combined video signal comprising:

processor means that process a presentation description and that control the manner in which images are combined;

memory means that store software executable by said processor means and that store video images;

tuner/decoder means that receive a first video signal and a second video signal and that route control information contained in said first video signal to said processor means and that route first video information from said first video

10 signal and second video information from said second video signal to decoder  
means;

decoder means that decode said first video information and produce a first  
video image in said memory means and that decode said second video  
information and produce a second video image in said memory means;

15 presentation description means that specify the manner in which a portion  
of said first video image is combined with a portion of said second video image to  
produce a combined image; and

video output means that output said combined image to a display device.